DOCUMENT RESUME

ED 454 872 IR 058 163

AUTHOR Calhoun, Karen
TITLE Redesign of Library Workflows: Experimental Models for Electronic Resource Description.
PUB DATE 2000-11-00
NOTE 16p.; In: Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web (Washington, DC, November 15-17, 2000); see IR 058 144.
AVAILABLE FROM Full text: http://lcweb.loc.gov/catdir/bibcontrol/calhoun_paper.html .
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Libraries; *Access to Information; *Cataloging; Higher Education; *Library Services; Models; Research Libraries
IDENTIFIERS *Electronic Resources

ABSTRACT

This paper explores the potential for and progress of a gradual transition from a highly centralized model for cataloging to an iterative, collaborative, and broadly distributed model for electronic resource description. The purpose is to alert library managers to some experiments underway and to help them conceptualize new methods for defining, planning, and leading the e-resource description process under moderate to severe time and staffing constraints. The first section compares catalogs and bibliographies. The second section discusses the influence of the Internet and licensed electronic resources, including working assumptions for bibliographic control and typical progress of a new electronic resource. The third section addresses accommodations in ARL (Association of Research Libraries) libraries, including discovery of sample networked resources in seven ARLs. The fourth section presents a new model for resource description. The fifth section describes the following selected experimental models: (1) CORC (Cooperative Online Resource Catalog) at Brown University (Rhode Island) and Cornell University (New York)--resource descriptions from collection development, reference, and cataloging; (2) the National Agricultural Library--resource descriptions from authors; (3) Yale University (Connecticut) and the record set for EBSCO Academic Search Elite--resource descriptions from vendors; and (4) the University of Tennessee-Knoxville and the Rochester Institute of Technology--resource descriptions from information technology staff. (Contains 24 references.) (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Redesign of Library Workflows: Experimental Models for Electronic Resource Description

ED 454 872

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

B. Wiggins

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Karen Calhoun

Director
Central Technical Services
Cornell University Library

Final version December 2000

Contents:

Looking Back: Technical Services as We've Known Them
Catalogs versus Bibliographies
A New Information Space: the Influence of the Internet and Licensed E-Resources
Space Walks: Heroic Accommodations in ARL Libraries
Towards New Models for Resource Description
Selected Experimental Models
Discussion
Looking Forward: Choosing Which Problems to Solve
Conclusion
Endnotes

Looking Back: Technical Services as We've Known Them

Ten years ago, Younger and Gapen described their vision of technical services in the year 2001. They predicted a paradigm shift characterized by a renewed focus on user needs, client-centered organizational structures, and the merging of technical and public services departments. (1) It hasn't happened. Instead, in academic libraries I know and have known, the most common organizational design remains a functional structure that uses the selection, acquisition, organization, and dissemination of library materials as the basis for logically grouping people and work. The main departments in academic research libraries continue to be collection development, technical services, and public services. However, in the past ten years an upstart department-library systems-has made the traditional library triad into a quadriad (Beile and Adams 2000). (2)

As noted by Stueart and Moran (1993) in their text on library management, a functional organizational structure has distinct advantages: it groups people and tasks that are similar, allows for specialization, and keeps library administrators keenly aware of the contributions and needs of each group. The functional organization's disadvantages include the competition that inevitably arises among departments; a focus on departmental rather than library-wide issues and goals; and difficulty collaborating across departments. (3)

Within technical services, those who organize and process materials for patron use-catalogers-both possess and take great pride in their in-depth knowledge of their specialization-resource description, better known as cataloging. Their chief product is the library catalog, and the process of building and maintaining it has been highly centralized within technical services departments. Library catalogs have served library staff and users well.

IR058163

The centralized technical services concept has not been without its critics. Younger and Gapen note that in centralized departments "there is too much attention directed toward library processing activities with insufficient focus and attention on meeting users' needs." (4) Noting that "our existing structures are no longer adequate to manage a digital or combination digital/traditional library," Boissonnas (2001) argues for "the deep integration of technical with other reader services" to prepare for the future and to overcome the "fragmentation, overspecialization, and philosophical inertia" associated with the functional division of library work. (5)

BEST COPY AVAILABLE

Boissonnas, Younger and Gapen's arguments are compelling, and they point to ways in which libraries must change. Nevertheless I maintain that the organizational structures and workflows that have brought us to this point in the development of present-day library catalogs have generally been efficient and effective. Tremendous waves of change have swept over cataloging departments-among them, the advent of shared online cataloging systems, the shift from card to online catalogs, strong growth in special or non-book collections, downsizing, outsourcing, and significant shifts in what professional catalog librarians are expected to do (or not to do). Catalog librarians and their managers-while they have not always done it with tremendous grace-have coped with every wave. That they have managed to do so is an important but largely unacknowledged miracle of the profession. Along the way, they have produced millions of catalog records for their own and the world's libraries, saving millions of dollars by reducing redundant effort by sharing their records in the bibliographic utilities.

The work to reap the full benefits of cooperative cataloging has continued in the past decade, resulting in new initiatives such as the Program for Cooperative Cataloging (PCC). Thomas and Younger (1993) spoke for many when they stated their vision for shared cataloging: "to put into place the necessary support to catalog, once and only once, every item owned or made accessible by libraries and to share that information with all others who need it." (6)

Today, despite the dramatic waves of change I've noted here, and in spite of the exhortations of visionaries, the basic approach to cataloging has not varied. Cataloging departments in academic libraries (and the people in them) remain much the same in terms of organizational structure, role and purpose in the library, understanding of the principles of the catalog, and professional values. Yes, new tasks and responsibilities have been taken on, and continual process improvements have been sought and made, but as overlays or add-ons to the same centralized service model as before. Outsourcing, for example, which provides an external source of catalog records, is by nature compatible with the kind of cataloging that would be done in-house. The PCC, while it represents an essential next step for cooperative cataloging as we have know it, continues to focus on a single set of library-centric resource description standards (AACR2 and MARC) and to not just rely on, but reinforce current library organizational structures. Surprisingly like the academic library reference service departments described by Ferguson (2001), cataloging "remains structured steadfastly around physical objects and the library as place." (7)

Our past experience of technical services in libraries is a powerful lens on how we see the present and the future. These experiences, together with the natural tendency to think the future is going to be like the past, can lead to such strong preconceptions about what resource description is, and who can and should do it, that we ignore critical facts that are in some way external to our mindsets. One such set of critical facts has to do with the history of the division between catalogs and bibliographies.

Catalogs versus Bibliographies

In his 1992 monograph on redesigning library services, Michael Buckland includes a chapter on bibliographies and catalogs.(8) Both bibliographies and catalogs contain resource descriptions, (9) albeit done according to varying standards and conventions, and both are forms of bibliographic control. Arguing that library cataloging can be viewed as a special case of bibliography, Buckland notes that the catalog is like a bibliography in that it is composed of information about works and editions of works. Unlike a bibliography, however, the catalog also concerns itself with individual copies of works-that is, the particular copies that a library holds. He goes on to point out the catalog's usual focus on one particular level of description-for monographs, the edition; and for serials, the title.

By contrast, bibliographies commonly list works at many levels of description (e.g., not only books and serials but also individual journal articles and conference papers). Bibliographies are generally the domain of not only reference librarians, but also individuals and groups operating outside librarianship, such as scholars, professional and scientific societies, government agencies and publishers. In particular, the role of large-scale indexing and abstracting of articles in periodicals has by tradition been left to publishers such as H.W. Wilson.

Buckland hastens to stress that the policy of excluding analytics from library catalogs is a matter of library tradition, not of principle. He concludes that the fact that the catalog is not normally thought of as bibliography "is largely an accident of semantic custom and of a tradition in library organization that associates the catalog with catalogers ... and bibliography with reference librarians."(10)

If one takes a long view of bibliographic control practices and history, then, the responsibility for resource description has been distributed among different groups inside and outside the library for a long time. Yet the boundaries between groups have been drawn so clearly, and the traditional arrangement has worked so well, that many librarians no longer recognize that the present division of labor is only one option among many for getting the work done.

A New Information Space: the Influence of the Internet and Licensed Electronic Resources

Academic libraries have no choice but to respond to the technology and applications of the Internet. Lubans' series of studies of student Internet use at Duke University has documented university students' "growing, even escalating use of Internet resources." (11) Students want the library to offer more Internet-based services. In my own research with Zsuzsa Koltay (Calhoun and Koltay 1999) into users' perceptions of the Cornell Library Gateway, I was struck by a student's comment: "The Gateway is the best [information system] I've ever used, but it is less than optimal. It is a great mock-up of the future." (12) Participants in the Cornell user study wanted the library to continue to add e-resources, especially full text; better communicate with them about the library's networked resources; provide multiple ways of discovering e-resources; and help them help themselves.

In 1992 Buckland proposed a redefinition of the catalog that foreshadowed the explosion of Internet resources. He argued for linking the information in online bibliographies with library holdings and permitting extended searches of multiple bibliographies and catalogs using multiple retrieval systems. (13) By 1999, Van de Sompel and Hochstenbach were experimenting with linking related information entities of all types-citation databases to catalogs and full text, finding aids to primary sources, catalog records to book reviews and images, and more. (14)

Clearly, the future of technical services is a future with Internet technology, applications, and resources in it; but from a technical services perspective, Internet resources tend to break the mold. They cause problems. They challenge our notions of the form and function of the catalog. The rules for cataloging them are in a nascent state. Our present exacting cataloging methods are too slow to handle their volume and complexity in reasonable turnaround times. They change so often that they overwhelm our capacity to maintain them. They force us to question conventional assumptions and workflows.

In keeping with the analysis laid out by Ercegovac (1997) (15), I suggest that Internet resources are driving fundamental changes that demand new operational and organizational assumptions about bibliographic control. The new assumptions are outlined in Table 1.

Table 1. Working Assumptions for Bibliographic Control

NOW	EMERGING
• Local collection, mostly print	• Many kinds of data sets, local and remote
• Catalog represents the collection but is separate from it	• With full text, catalog and collection are converging
• Highly standardized bibliographic records in library schemes (AACR2, MARC format, LC or DDC class, LCSH, Sears)	• Less structure in indexing, mixed representations of data; metadata can be prescribed by varying rules or be free form
• Centralized responsibility for resource description/metadata creation and limited decentralization for specific subjects or languages	• Highly distributed responsibility for resource description/metadata creation; records come from multiple sources

Operationally, electronic resources drive the catalog away from bibliographic control of a physical collection toward the representation and control of a virtual repository and the possibility of a new catalog as described by Buckland. (16) Organizationally, the proliferation of Internet resources is causing a technical services identity crisis.

Most libraries' technical services departments reflect policies and practices that are outgrowths of functional organizational structures. Technical services departments tend to be staffed with individuals who are hired and trained to be experts in some aspect of the

acquisition or organization of library materials. As is typical of experts within any type of organization (Neuhauser 1988), it is not unusual for these individuals to be tightly focused on the tasks they perform, to have minimal contacts outside technical services, and to be unfamiliar with the library activities outside their particular function. (17)

This organizational structure has worked well for print resources-selecting, acquiring and describing them is nearly always a linear, sequential process in which one person or group works independently on each step. This is possible because policies and rules are well established and known to participants. But, as can be seen from the workflow description in Table 2, selecting, "acquiring," and describing Internet resources, can be (and often is) an iterative, highly collaborative, looping process that can involve many individuals from many functional groups. The outputs of the process can also vary-the end result may be MARC records in the catalog, links and summary descriptions on a library Web page, or even records in a non-MARC metadata format (Dublin Core or some locally developed record format).

Table 2. "Typical" Progress of a New Electronic Resource

	Steps
1	A selector identifies and selects an electronic resource
2	The selector initiates a request to acquire and/or describe the resource (and/or list it on one or more library Web pages)
3	Acquisitions/selectors/information technology/catalogers/reference staff exchange inquiries as needed
4	Acquisitions or collection development or reference staff negotiate with vendor/publisher/author (for licensed resources)
5	Acquisitions initiates request to describe the resource and/or add it to appropriate Web list
6	Acquisitions/catalogers/selectors/information technology/reference staff exchange inquiries as needed
7	Catalogers consult the resource, resource description standards and databases to prepare resource descriptions (however in some academic libraries it is more common for a resource description to be added to one or more Web lists or databases than for it to be cataloged)
8	Another round of inquiries as needed
9	The resource description is added to the catalog, and/or it gets listed on one or more library Web pages, and/or included in a locally-created searchable database of electronic resources

Space Walks: Heroic Accommodations in ARL Libraries

There are countless variations on the "typical" process sketched in Table 2, because libraries have chosen varied models for organizing themselves to manage electronic resources. In an attempt to get a clearer picture of how libraries are accommodating Internet resources now, I completed a brief analysis of the resource descriptions provided for a set of seventeen commonly-licensed online databases and full text journals at seven of the largest ARL libraries in the United States. (18) I did not include titles of e-books in my analysis,

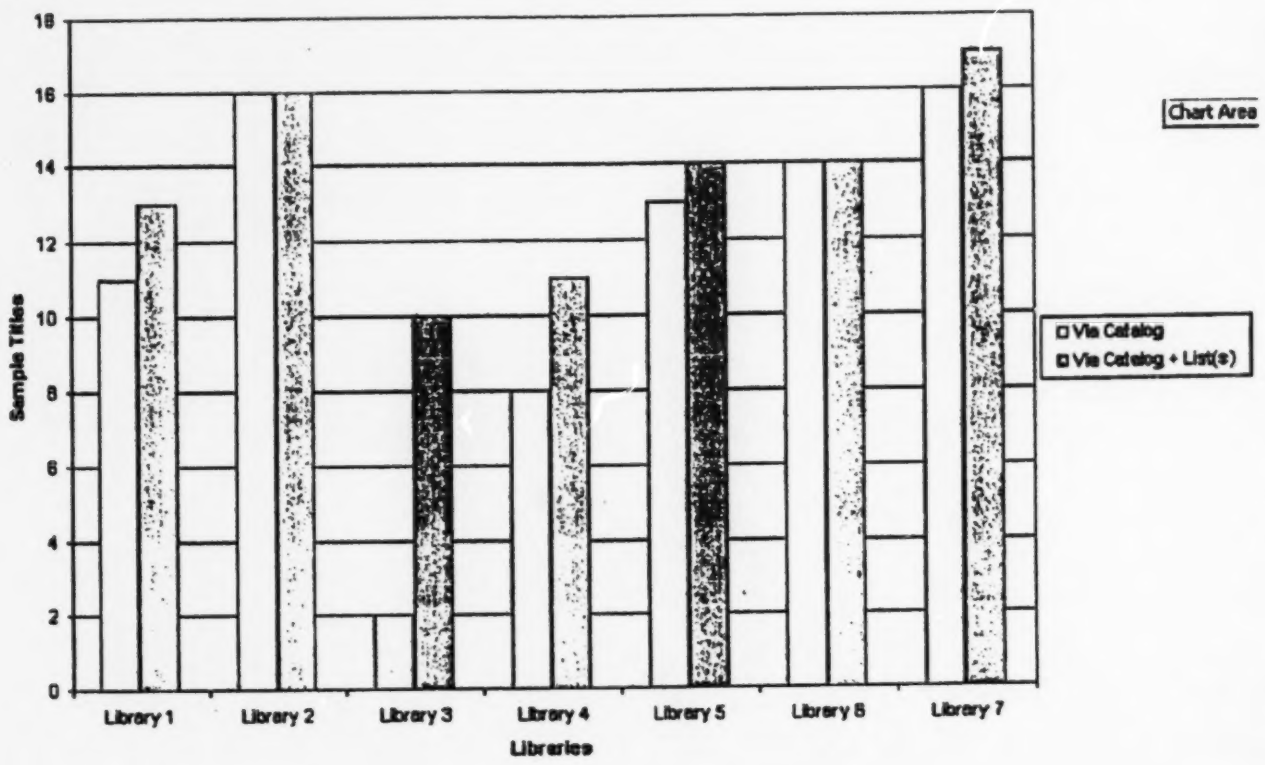
although they are approaching us so rapidly now, because how to provide resource descriptions for e-books is still under intense discussion in libraries.

There are two models of networked resource description in operation among the ARL libraries examined. All of the sample libraries provide for discovery and access of online databases and electronic full text journals both via the catalog and via the library Web site. In many cases it appears that collection development or public services is responsible for mounting and maintaining the Web list or lists of titles, while records in the catalog continue to be the domain of technical services.

The format and number of the Web lists vary widely from library to library. One library maintains a searchable database of its networked databases and full text journals that serves as a single point of entry; there are no separate lists, although the contents of the searchable database may be browsed by broad subject category. Two other libraries maintain searchable lists; one is a searchable list of e-journals (no databases) and the other is a searchable list of databases (no e-journals). All of the others maintain lists of networked resources that can be browsed either alphabetically or by broad subject categories. In most cases there are two lists, one for databases and one for e-journals, but in a few cases there are multiple lists of both, and one needs to know which list to pick (e.g., the science and technology one or the social sciences one) before beginning to browse for a particular title.

Figure 1 provides a graphical view of the number of titles that I was able to discover via the catalog alone, or by a combination of using the catalog and available Web lists, for the seven libraries examined. Using the catalog alone, I was able to discover about nine titles, or 54%, on average. (19) The minimum number of titles discovered via the catalog alone was two, or about 12%; the maximum sixteen, or 94%. Adding in what can be found using both the catalog and Web lists, the average number of titles discovered rises to about fourteen (82%), the minimum to eleven (65%), and the maximum to seventeen (100%).

Figure 1. Discovery of Sample Networked Resources in Seven ARLs



BEST COPY AVAILABLE

The findings provide strong evidence that different functional groups in ARL libraries are already distributing the work of creating resource descriptions for electronic resources. These librarians are making heroic accommodations for Internet resources and adapting their methods to include them. Nevertheless, the traditional boundaries among the functional groups appear to remain intact, although sometimes blurry; few or no libraries appear to be starting over with new service models or examining basic assumptions of specialization. I base this conclusion on my own experiences working with library Web sites and catalogs, conversations with colleagues, and the specific findings of this analysis. Library users probably do want and need multiple ways to discover networked resources. Yet the reality suggested by this analysis is that searchers often must use both the catalog and library Web lists to discover what the library makes available, and even when they do use both, they do not always get the full picture.

This analysis and my own experience further suggest that libraries' current methods for producing electronic resource descriptions generally result in suboptimal, fragmented discovery and retrieval systems that are difficult for library patrons to understand and use. The double work that library staff are often expending to provide multiple access methods-via the catalog and lists-is generally is not paying off as it could, because the efforts by different functional groups tend to be uncoordinated and poorly integrated.

At a minimum, we must redesign and integrate the functionality of our libraries' catalogs and Web sites so they can function as a coherent information system. Accomplishing this is difficult, because libraries' heroic accommodations to date are for the most part overlays and add-ons to the same operational and organizational models as before. We can do better, but to build truly coherent, usable and useful library systems that successfully integrate networked resources with our collections and services, we must be willing to transform ourselves, our methods, and our conventional organizational structures.

Toward New Models for Resource Description

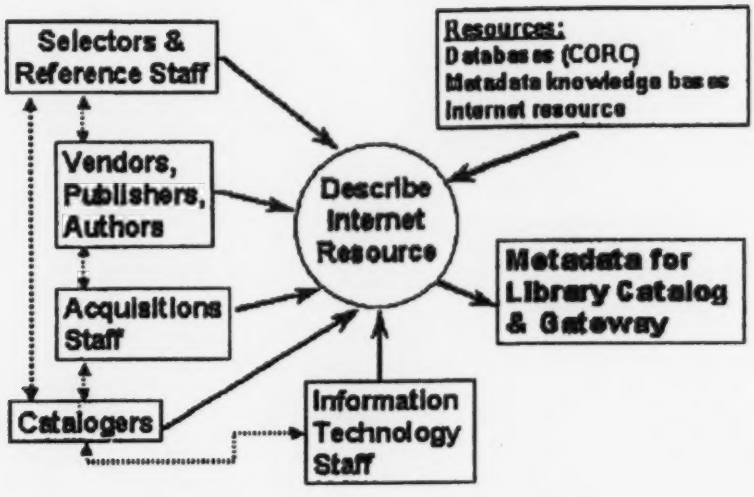
The call to fundamentally change organizational processes in order to achieve quantum leaps in organizational effectiveness is not new. In their pioneering book Reengineering the Corporation, Michael Hammer and James Champy (1993) call for "putting back together again the work that Adam Smith and Henry Ford broke into tiny pieces so many years ago." (20) They argue for "process teams" or "virtual teams" that obviate the need for hand-offs among functional departments and thereby dramatically speed up operations. These kind of teams are responsible not for a fragment of a process, but for an entire cross-functional process, such as handling a customer complaint, or completing the design of a new product or service. (21)

Along these same lines, I am proposing that building effective, coherent systems for the discovery and retrieval of library-selected resources will require us to brush aside conventional wisdom and start over. The first step is to put aside for the moment the assumptions we have made about the proper division of labor for producing resource descriptions. The second step is to step back from what we know about the appropriate level of description in a catalog and build systems that will allow library users to discover and navigate freely among resource descriptions and an array of heterogeneous collections. The third step is to rethink and expand our notions of standardization and controlled vocabulary.

How might we begin to reconceptualize organizational structures and processes for electronic resources? As demonstrated by Richardson (1999) in his modeling of the reference transaction, the tools of systems analysis provide for a graphic representation of inputs and outputs and a top-down perspective that can lead to new insights about a process. (22) The following section of this paper is an adaptation of Richardson's systems analysis approach; it attempts to model the electronic resource description process, and the players in it, as a system.

Figure 2 depicts the electronic resource description process as a widely distributed one, with many players inside and outside the library. It is intended to be an illustrative rather than comprehensive illustration of the process. Solid lines designate the flows of data in the system. The dotted line flows designate inquiries and responses exchanged among the various players (for the sake of clarity, some the inquiry/response lines have been omitted from the figure, but the intention is that inquiries and responses would flow between all players). The principal output of this system is metadata, or resource descriptions. It is assumed that the metadata provides a link to the resource itself.

Figure 2. Context of a Widely Distributed Electronic Resource Description Process



In this proposed system, all functional groups in the library could conceivably contribute resource descriptions. Resource descriptions could also come from vendors, publishers, or authors themselves. Data will also flow in from resource description databases (e.g., CORC), from metadata knowledge bases (e.g., Dublin Core, EAD, and MARC/AACR2), and from the Internet resources themselves. The process ends with the output of metadata and its integration into the library information discovery and retrieval system. However, recognizing that the networked resource description process can be dynamic, the initial content of a resource description could be quite minimal, but it could then be modified and enriched over time.

As far as I know a fully-realized version of the system does not exist anywhere. But the proposed system can serve as a model for envisioning the context and data flows of future resource description processes. In addition, the model can help us respond creatively to the changes we face in libraries by freeing our minds from the past, stepping back from our mindsets, and allowing us to see what we need to do with fresh eyes. The following sections of this paper contain descriptions of various innovations at a number of libraries that have realized at least parts of the model of a widely distributed e-resource description process.

Selected Experimental Models

CORC at Brown and Cornell: Resource Descriptions from Collection Development, Reference and Cataloging

CORC is a cooperative effort to create a library-selected database of Web-based electronic resource descriptions. It is hosted by OCLC. (23) Prior to becoming production system in July 2000, CORC was a research initiative of several hundred libraries collaborating with the OCLC Office of Research. Some CORC participants have used the system as a framework for collaboration among different functional groups in their libraries. Two of these are Brown and Cornell.

The research phase of CORC became available at a time that librarians at Brown were discussing how to improve the information they provided to users to alert them to quality Web sites. They were interested in producing both Web-based bibliographies and cataloging records. Brown reference and collection development librarians used the fifteen elements of Dublin Core as the first step toward producing a MARC resource description for a Web resource. Catalogers then finished them in MARC and exported them to the local

BEST COPY AVAILABLE

catalog. In addition, Brown subject specialists and reference librarians designed or customized pathfinders based on existing sites cataloged in the CORC database. In their article on CORC at Brown, Caldwell, Coulombe, Fark and Jackson (2001) express their satisfaction with the project outcomes, not the least of which was an enhancement of shared values among staff from different functional groups. They remarked, "for too long, it has been assumed that public and technical services cannot work closely together in a cataloging project because of differing agendas, missions, and skills. We at the Brown University Library have shown that it is possible and fruitful for both." (24)

The CORC project undertaken at Cornell was deliberately structured as a research project. We sought to take advantage of CORC's simultaneous support for Dublin Core and MARC as a breakthrough technology that would enable us to experiment, at low levels of risk and cost, with broadly distributed resource description.

For the duration of our experiment, we were a virtual team of three bibliographers, a reference librarian, two catalogers, and a project coordinator. The bibliographers and reference librarian selected the desired Internet resources and began the resource descriptions using Dublin Core. They supplied at least a title, URL, and a summary description. Next, the catalogers retrieved the records from the Cornell in-process file, finished them in MARC format, and then exported them to our local catalog and Gateway. While there was a hand-off of work from collection development/reference to cataloging, it was not the same as hand-offs between departments that are not organizationally integrated who in fact operate in functional "silos." Instead, the group functioned as a close-knit team, each member of which brought a different set of skills and perspectives to the work. Our workflow sought to put back together again a group that has been artificially separated by organization.

Cornell's experiences and results from the CORC project are reported elsewhere (Calhoun et al. 1999). (25) In summary, we found that the changes we made to the traditional workflow can ease and streamline the production of Internet resource descriptions. We found that distributed resource description is both feasible and beneficial, and that bibliographers and reference librarians can readily use Dublin Core to create preliminary records using CORC.

Another major plus was finding records already in CORC for most of the resources that were selected during the project. CORC is not only an extremely important advance in the library community's cooperative cataloging model, but it has the potential to expand the benefits of cooperation to new communities that need resource descriptions. Along the same lines, the team agreed that the most productive conversations about DC and MARC would assume that both have their place at Cornell. We should focus on how to forge a complementary relationship between the two standards, striving to optimize the strengths of each.

National Agricultural Library: Resource Descriptions from Authors

At the recent ALCTS Directors of Technical Services in Large Research Libraries meeting in Chicago, Sally Sinn of the National Agricultural Library (NAL) reported that scientists are creating and submitting resource descriptions of their work, and NAL librarians are working closely with them. There are two collaborations underway that are part of a redesign of NAL's RMIS (Research Management Information System) initiated by the Agricultural Research Service. One project is an effort that might be described as "indexing-in-publication" in which scientists submit descriptions for potential and completed publications (articles, book chapters, and conference papers) that is then standardized to match AGRICOLA citation format. In another project, NAL is developing a thesaurus of hierarchically arranged topics to describe the scientists' research projects for improved subject retrieval using standard vocabulary. (26)

Yale and the Record Set for EBSCO Academic Search Elite: Resource Descriptions from Vendors>

Late in 1998, the Program for Cooperative Cataloging's Standing Committee on Automation convened a Task Group on Journals in Aggregator Databases to (1) propose the content of vendor-supplied records for the full text journals in aggregator databases like ProQuest and (2) complete a demonstration project with an interested vendor. The initial task group completed its charge and issued a final report (Riemer and Calhoun 2000) (27), then was reconstituted for two more years to continue lobbying vendors to create record sets for their products and to pursue new areas of research.

In early 1999, EBSCO developers began collaborating with the task group to produce a record set for the approximately 1,100 titles of full text journals accessible from Academic Search Elite. The records, which are derived from CONSER records for the corresponding print journals, have been available for download by EBSCO customers at no charge since summer 1999. EBSCO periodically reissues the set to reflect additions, deletions and changes to their product.

Several libraries have acquired the record set from EBSCO and loaded them into their catalogs. One of the libraries was Yale. Matthew

Beacom, Catalog Librarian for Networked Information Resources at Yale, reported that the load generally went well.

Yale has also loaded the first set of updates to the initial load well as the records for EBSCO Business Source Premier titles. (28) By doing so, Yale has greatly enhanced its users' ability to discover via the catalog what full text resources are available to them, yet with a minimum of effort expended by its technical services and information technology departments.

The University of Tennessee-Knoxville and the Rochester Institute of Technology: Resource Descriptions from Information Technology Staff

A system of widely distributed resource description opens the door to broader participation of the library's information technology staff. Resource descriptions can be produced automatically, for example for the full text titles of e-journals in a vendor or publisher's database. Britten and others (2000) reported they harvested data about full text titles from vendors' Web sites and subsequently massaged them with Perl scripts and a utility called MarcMakr. (29) The end product was a set of MARC resource descriptions for the full text journals in several large aggregator databases for the catalog at the University of Tennessee at Knoxville.

At the Rochester Institute of Technology (RIT), a library wide task force was charged with finding an inexpensive solution to placing as many e-journals as possible under bibliographic control. Jiras (2000) reported that E-journal Web pages were becoming difficult to organize, providing full cataloging treatment for each title was too slow and labor intensive, and trying to keep up with added and cancelled titles and changes in holdings was a losing battle. The initiative that grew out of the task force's work led to a process in which library systems staff produced resource descriptions for aggregator e-journal titles by harvesting data from the vendor sites, massaging the data in several ways to produce MARC records, then loading the records into the catalog. (30)

Discussion

The library quadriad of collection development, public services, technical services and library systems is a persistent and highly successful organizational model. It is a functional division of labor that has the advantage of allowing specialization; in particular, catalogers' ability to focus on the catalog has produced millions of records for their own and the world's libraries in the past thirty years. Nevertheless, the functional division of labor has the disadvantages of fragmenting library processes, making cross-functional collaboration difficult, and discouraging "out of the box" thinking.

The incredible demand for Internet resources gives libraries strong incentives to reunite and intelligently coordinate the efforts of the individuals and groups that have always shared the work of resource description. In fact, an uneasy collaboration of cataloging, collection development, and reference librarians is already in evidence. There is already a two-pronged approach to the discovery and retrieval of electronic resources by users: catalog records and Web lists. But the two-pronged approach is too often uncoordinated and less than library users deserve. To do better, libraries must reintegrate the process of, and responsibility for, electronic resource description. Doing so is an essential first step in building a coherent, usable and useful library information discovery and retrieval system.

This paper proposes an electronic resource description process that could be very effective for making resource descriptions available more quickly, in greater numbers, and at less cost, assuming the process delivers metadata that is useful to readers. Lundgren and Simpson (1999) have explored the question of what is useful metadata. (31) I have touched on only a few of the experiments that are underway now to broaden and integrate participation in libraries' electronic resource description processes.

Looking Forward: Choosing Which Problems to Solve

Given the inevitable constraints on human, financial, and temporal resources, it is critical that librarians focus their energies on solving problems that will help their organizations in the future. As Buckland argued early in the decade, it is time to redefine the catalog, stop wasting effort on outdated models, and adopt a new bibliographic strategy. He urges librarians to think in terms of making use of all networked bibliographies and catalogs, not just local ones, and he lays out a set of basic functional requirements for a more universal approach to library collections. Similarly, in explicating their concept of "reference linking," Van de Sompel and Hochstenbach urge an evolution toward connecting all the available information, in order to come to a fully interlinked information environment. (32)

What might be some of the functional requirements of the information discovery and retrieval system proposed in this paper? One requirement would certainly be to deliver resource descriptions that are useful to readers. This would require us to answer the questions

BEST COPY AVAILABLE

of what is a useful resource description, from a user's perspective, and how the needs of various user groups differ. As mentioned previously, Lundgren and Simpson have begun work on these questions, and I encourage others to take up where they left off.

Another requirement would be to support discovery and retrieval of a resource that is described at a full range of levels of granularity (e.g., citations and full text of articles, books, serials, sound recordings, images, the content of digital collections). A system that provides access to an array of information resources, both print and electronic, must provide contextual information and guidance to help users make sense of the results of their searches. (33)

A third system requirement would be a supporting infrastructure that brings order from the chaos inherent in this loose federation of data from many sources. Vellucci (1997) calls this a "metacatalog." It will contain resource descriptions in multiple metadata formats, created according to multiple standards, with name and subject headings created according to different communities' conventions, yet its infrastructure must present search results to searchers in a sensible way. On this point Vellucci says "the next generation metacatalogs should be able to access all relevant information seamlessly ... In order to accomplish this, each stakeholder community must ... concentrate on developing ways to layer, exchange and translate data within a loosely-coupled organizational system." (34)

The critical need for systems to be able to manage loosely federated data from many sources is far from unique to libraries, and it is not new. In 1998 the National Science Foundation, the Biological Resources Division of the U.S. Geological Survey, and an ALCTS task force hosted a Taxonomic Authority Files (TAF) Workshop. (35) The purpose was to bring together members of the biological sciences and library communities to explore the highly partitioned information environment in the biological sciences, to describe authority control in libraries, and to discuss the possibilities for managing widely distributed biological data sources to achieve consistency across shared concepts and names.

At the TAF workshop I presented an overview of authority control in libraries and concluded my talk with a call for a number of improvements to library authority control. Among these suggestions were to abandon the notion of a single, monolithic, all-encompassing global authority file in favor of a system of linked interoperable files; to deeply integrate authority data into end-user information systems (e.g., mapping a searcher's query into the vocabulary or naming conventions of the database being searched); and to better integrate the library community's authority control conventions with those of the abstracting and indexing community. I also noted that "taking any significant action would surely require a rethinking of the library community's cu... model of authority control." (36)

At the same workshop Stuart Nelson of the National Library of Medicine described the multi-thes...rus system called UMLS (Unified Medical Language System). (37) At the beginning of his talk Nelson used a Biblical analogy to illustrate the problems of a diverse, complex information system: "much of what the UMLS is approaching is ... [the problem] depicted in the story of the Tower of Babel." The purpose of UMLS is to retrieve and integrate information from patient records, databanks, bibliographic databases, full text sources, and elsewhere. An integral piece of the UMLS is a metathesaurus that includes data about naming conventions in a variety of different systems using a variety of controlled vocabularies.

These are the kinds of problems that librarians who are building the next generation of information discovery and retrieval systems will need to grapple with and solve.

Conclusion

Libraries appear willing to experiment, but I anticipate many obstacles to the full deployment of a widely distributed electronic resource description process. I believe the principal obstacles will not be technical or operational, but organizational and attitudinal. Many librarians are deeply vested in existing processes and organizational structures. Not only that, the existing processes and structures still function well for most items that are added to the collections. Finally, because we are and will be in a transitional state for some time, librarians must strike an appropriate balance between their everyday work and new ways of doing things.

Perhaps in the near term, then, it would be more practical for libraries to avoid radical restructurings and instead make liberal and frequent use of virtual teams. These are cross-functional groups that exist alongside (and sometimes outside) the formal organizational structure. For example, for the purpose of introducing a new electronic resource description process, or providing ongoing support for one, the members of the virtual team would share authority and accountability. At the same time, the members would continue to report to different individuals and departments in the library hierarchy.

Beyond the redesign of library workflows, building the coherent, usable and useful information discovery and retrieval system I have proposed will require determination, perseverance, and skills from all walks of librarianship. I am convinced that technical services

librarians have a great deal to bring to the table, provided they tap into their creativity and apply their significant knowledge of bibliographic control to the new information space in which libraries operate.

---

1. Younger, Jennifer A. and D. Kaye Gapen. 1990. Technical services organization: where we have been and where we are going. In Technical services today and tomorrow, ed. Michael Gorman. 171-83. Englewood CO: Libraries Unlimited.

2. Beile, Penny M. and Megan M. Adams. 2000. Other duties as assigned: emerging trends in the academic library job market. College & research libraries 61, no. 4: 336-47.

3. Stueart, Robert D. and Barbara B. Moran. 1993. Library and information center management. 4th ed. Englewood CO: Libraries Unlimited, 80-1.

4. Younger and Gapen, Technical services organization, 176.

5. Boissonnas, Christian M. 2001. Technical services: the other reader service. portal: libraries and the academy 1. In press.

6. Thomas, Sarah E. and Jennifer A. Younger. 1993. Cooperative cataloging: a vision for the future. Cataloging & classification quarterly 17, nos. 3, 4: 257.

7. Ferguson, Chris. 2000. "Shaking the conceptual foundations," too: integrating research and technology support for the next generation of information service. College & research libraries 61, no. 4: 300-11. The title within the article title refers to Jerry Campbell's 1992 article in Reference services review. While Ferguson is discussing the dramatic changes in reference services and calling for dramatic new approaches, the parallels with the history and current environment in technical services are striking. I particularly appreciated Ferguson's insight into the difference between "layering" new services upon old models, versus creating new service models, as well as his use of the phrase "heroic accommodations," which I have borrowed for this article.

8. Buckland, Michael. 1992. Bibliographic access reconsidered. Ch. 4 of Redesigning library services: a manifesto. 24-41. Chicago: American Library Association.

9. For the purpose of this paper, resource description is defined broadly to include the creation of any surrogate of an item whose purpose is to facilitate library users' discovery and retrieval of the item.

10. Buckland, Bibliographic access reconsidered, 29.

11. Lubans, John. 1999. Students & the Internet: spring 1999 survey (study 3). Available at: http://www.lib.duke.edu/lubans/docs/study3.html. Accessed: August 11, 2000.

12. Calhoun, Karen and Zsuzsa Koltay. 1999. Library gateway focus groups report, January 1999. Available at: http://www.library.cornell.edu/staffweb/GateEval/contents.html. Accessed: August 11, 2000.

13. Buckland, Redesigning library services, 32-9.

14. Van de Sompel, Herbert and Patrick Hochstenbach. 1999. Reference linking in a hybrid library environment, parts 1 and 2. D-Lib magazine 5, no. 4 (April). Available at: http://www.dlib.org/dlib/april99/04contents.html. Accessed: August 11, 2000.

15. Ercegovac, Zorana. 1997. The interpretations of library use in the age of digital libraries: virtualizing the name. Library & information science research 19, no. 1: 35-51. Table 1 is an adaptation of Ercegovac's table "Toward digital libraries" on page 42.

16. Buckland, Redesigning library services, 32-

17. Neuhauser, Peg. 1988. Characteristics of organizational tribes. Ch. 2 of Tribal warfare in organizations. 15. Cambridge MA: Ballinger Publishing.

18. I chose seventeen titles to look for--ten online databases and seven full text journals. The list below provides the titles and the sources from which they are available. I selected them because (1) they are commonly licensed by large ARLs and thus likely to be accessible to their users; and (2) they present resource description challenges of different types and levels (simple to complex). For example, Callaloo is available in JSTOR, a full text collection with a stable list of titles that are well maintained over time, while American Heritage is part of several large, amorphous vendor aggregations with shifting sets of titles. The seven libraries were Harvard, Yale, UCLA, the University of Illinois at Urbana, the University of Michigan, Columbia, and Cornell.

Titles Examined	Source(s)
ABI Inform	Ovid, OCLC, Proquest
Avery index	RLG
Arts & humanities citation index	ISI Web of Science
Congressional Universe	CIS
Dissertation abstracts	UMI
ERIC	ERIC, Ovid, OCLC
INSPEC	IEE, Ovid
Academic Universe	Lexis Nexis
UnCover	CARL
WorldCat	OCLC
Harvard business review	Full text in Ovid ABI Inform, Proquest ABI Inform, EBSCO Academic Search Elite, etc.
Time	Full text in Ovid ABI Inform, Proquest ABI Inform, EBSCO Academic Search Elite, OCLC Periodical Abstracts, etc.
Callaloo	Full text in JSTOR, Project Muse, OCLC ECO, etc.
American Heritage	Full text in EBSCO Academic Search Elite, OCLC Periodical Abstracts, Proquest Periodical Abstracts, etc.
SIAM journal on applied mathematics	Full text in JSTOR, SIAM Journals Online, etc.
Wall Street Journal	Full text in Proquest Periodical Abstracts Research II, OCLC
Algorithmica	Full text in Springer Link, EBSCO Online, etc.

19. There were cases in which the library did not license the title in question, but these cases were rare. Usually the missing title was licensed, just not findable via the library's catalog or Web site.

20. Hammer, Michael and James Champy. 1993. The new world of work. Ch. 4 of Reengineering the corporation: a manifesto for business revolution. 65. New York: HarperBusiness.

21. Ibid., 66-7.

22. Richardson, John V. Jr. 1999. Understanding the reference transaction: a systems analysis perspective. College & research libraries 60, no. 3: 211-22.

23. For more information about CORC visit http://www.oclc.org/oclc/corc/index.htm. Accessed: August 11, 2000.

24. Caldwell, Ann Dominique Coulombe, Ronald Fark, and Michael Jackson. 2000. Never the twain shall meet? Collaboration between catalogers and reference librarians in the OCLC CORC project at Brown University. Journal of Internet cataloging 4, no. 1. In press.

25. Calhoun, Karen, et al. 1999. CORC at Cornell project: final report. Available at: http://campusgw.library.cornell.edu/corc/. Accessed: August 11, 2000.

26. Sinn, Sally. 2000. Reported in the minutes of the ALCTS Directors of Technical Services in Large Research Libraries meeting, July 7, 2000, and confirmed in e-mail exchange with the author, August 8, 2000.

27. Riemer, John and Karen Calhoun. 2000. PCC Standing Committee on Automation (SCA) Task Group on Journals in Aggregator Databases: final report, January 2000. Available at: http://lcweb.loc.gov/catdir/pcc/aggfinal.html. Accessed: August 11, 2000.

28. Beacom, Matthew. 2000. E-mail exchange with the author, August 8, 2000.

29. Britten, William A., et al. 2000. Access to periodicals holdings information: creating links between databases and the library catalog. Library collections, acquisitions and technical services 24, no. 1. In press.

30. Jiras, Jonathan. 2000. Access to e-journals at RIT. Presentation at Partners in Information and Innovation meeting, February 2, 2000, at Rensselaer Polytechnic Institute, Troy, NY.

31. Lundgren, Jimmie and Betsy Simpson. 1999. Looking through users' eyes: what do graduate students need to know about Internet resources via the library catalog? Journal of Internet cataloging 1, no. 4: 31-44.

32. Van de Sompel and Hochstenbach, Reference linking, part 1.

33. For example, in a fully interlinked information environment, a user's search for "George Washington" could retrieve images, correspondence and other primary source materials, books, articles in journals, audiovisual materials, etc. How can we provide enough information about where the user's "hits" are coming from to allow him or her to make sense of what has been retrieved and to navigate to what is wanted?

34. Vellucci, Sherry V. 1997. Options for organizing electronic resources: the coexistence of metadata. Bulletin of the American Society for Information Science 24, no. 1 (Oct./Nov.).

35. A description of the TAF workshop and the participants' papers may be found at http://research.calacademy.org/taf/proceedings/Proceedings.html. Accessed: August 11, 2000.

36. Calhoun, Karen. 1998. A bird's eye view of authority control in cataloging. Paper presented at TAF Workshop, June 22-23, 1998, in Washington DC. Available at: http://research.calacademy.org/taf/proceedings/Calhoun.html. Accessed: August 11, 2000.

37. Nelson, Stuart. 1998. The Unified Medical Language System: applicable experiences and observations. Paper presented at TAF Workshop, June 22-23, 1998, in Washington DC. Available at: http://research.calacademy.org/taf/proceedings/nelson/index.htm. Accessed: August 11, 2000.

Library of Congress
December 14, 2000
Comments: lcweb@loc.gov

Conference Home Page

What's new

Greetings from the Director for Cataloging

Topical discussion groups

NAS study and 2 articles from the LC staff *Gazette*

Conference program

Speakers, commentators, and papers

Conference sponsors

Conference discussion list

Logistical information for conference participants

Conference Organizing Team

Bicentennial Conference on Bibliographic Control for the New Millennium

Confronting the Challenges of Networked Resources and the Web

sponsored by the Library of Congress Cataloging Directorate

Karen Calhoun
Director, Central Technical Services
107-D Olin Library
Cornell University
Ithaca, NY 14853

Redesign of Library Workflows: Experimental Models for Electronic Resource Description

About the presenter:

Karen Calhoun, M.S., M.B.A., is the Director of Central Technical Services at Cornell University Library, a position she has held since March 2000. Prior to that she was the head of cataloging. Active in the development of Cornell's Library Gateway (http://campusgw.library.cornell.edu/) and a frequent speaker on technical services in the digital library, Karen's recent research and operational interests have focused on the organization of networked resources and services, user needs, project management, library workflows, cross-functional teams, and cooperative cataloging and authority control. Currently she leads Cornell's participation in the CORC project, chairs the Program for Cooperative Cataloging (PCC) Standing Committee on Automation and the ALCTS CCS Policy and Research Committee, is active in the PCC Task Group on Journals in Aggregator Databases, and serves as assistant editor of Library Collections, Acquisitions and Technical Services. In addition she has co-edited a special issue of the Journal of Internet Cataloging (forthcoming) on CORC. Before coming to Cornell she held positions at OCLC and the University of Oregon.

Full text of paper is available

BEST COPY AVAILABLE

Cataloging Directorate Home Page

Library of Congress Home Page

Summary:

This paper explores the potential for and progress of a gradual transition from a highly centralized model for cataloging to an iterative, collaborative, and broadly distributed model for electronic resource description. The author's purpose is to alert library managers to some experiments underway and to help them conceptualize new methods for defining, planning, and leading the e-resource description process under moderate to severe time and staffing constraints. To build a coherent library system for discovery and retrieval of networked resources, librarians and technologists are experimenting with team-based efforts and new workflows for metadata creation. In an emerging new service model for e-resource description, metadata can come from selectors, public service librarians, information technology staff, authors, vendors, publishers, and catalogers. Arguing that e-resource description demands a level of cross-functional collaboration and creative problem-solving that is often constrained by libraries' functional organizational structures, the author calls for reuniting functional groups into virtual teams that can integrate the e-resource description process, speed up operations, and provide better service. The paper includes an examination of the traditional division of labor for producing catalogs and bibliographies, a discussion of experiments that deploy a widely distributed e-resource description process (e.g., the use of CORC at Cornell and Brown), and an exploration of the results of a brief study of selected ARL libraries' e-resource discovery systems.

Library of Congress
December 14, 2000
Comments: lcweb@loc.gov

BEST COPY AVAILABLE